

CLAIMS

What is claimed is:

1. An isolated nucleic acid comprising a nucleic acid selected from the group consisting of
5 a nucleic acid encoding any of C-1027 open reading frames (ORFs) -7 through 42, excluding ORF 9 (cagA);
a nucleic acid encoding a polypeptide encoded by any of C-1027 open reading frames (ORFs) -7 through 42, excluding ORF 9 (cagA); and
a nucleic acid amplified by polymerase chain reaction (PCR) using
10 primer pairs that amplify any of C-1027 open reading frames (ORFs) -7 through 42, excluding ORF 9 (cagA).
2. The isolated nucleic acid of claim 1, wherein said nucleic acid comprises a nucleic acid encoding at least two open reading frames (ORFs) selected from the group consisting of ORF-1 through ORF 42, excluding ORF 9 (cagA).
- 15 3. The isolated nucleic acid of claim 1, wherein said nucleic acid comprises a nucleic acid encoding at least three open reading frames (ORFs) selected from the group consisting of ORF-1 through ORF 42, excluding ORF 9 (cagA).
4. An isolated nucleic acid comprising a nucleic acid that specifically hybridizes under stringent conditions to an open reading frame (ORF) of the C-1027
20 biosynthesis gene cluster, excluding ORF 9 (cagA), and can substitute for the ORF to which it specifically hybridizes to direct the synthesis of an enediyne.
5. The isolated nucleic acid of claim 4, wherein said isolated nucleic acid comprises a nucleic acid that specifically hybridizes under stringent conditions to a nucleic acid selected from the group consisting of ORF -7, ORF -6, ORF -5, ORF -4, ORF -3, ORF -
25 2, ORF -1, ORF 0, ORF 1, ORF 2, ORF 3, ORF 4, ORF 5, ORF 6, ORF 7, ORF 8, ORF 10, ORF 11, ORF 12, ORF 13, and ORF 14.
6. The isolated nucleic acid of claim 4, wherein said isolated nucleic acid comprises a nucleic acid that specifically hybridizes under stringent conditions to a nucleic

acid selected from the group consisting of ORF 15, ORF 16, ORF 17, ORF 18, ORF 19, ORF 20, ORF 21, ORF 22, ORF 23, ORF 24, ORF 25, ORF 26, ORF 27, ORF 28, ORF 29, ORF 30, ORF 31, ORF 32, ORF 33, ORF 34, ORF 35, ORF 36, ORF 37, ORF 38, ORF 39, ORF 40, ORF 41, and ORF 42.

5 7. The isolated nucleic acid of claim 5, wherein said isolated nucleic acid comprises a nucleic acid selected from the group consisting of ORF -7, ORF -6, ORF -5, ORF -4, ORF -3, ORF -2, ORF -1, ORF 0, ORF 1, ORF 2, ORF 3, ORF 4, ORF 5, ORF 6, ORF 7, ORF 8, ORF 10, ORF 11, ORF 12, ORF 13, and ORF 14.

10 8. The isolated nucleic acid of claim 6, wherein said isolated nucleic acid comprises a nucleic acid selected from the group consisting of ORF 15, ORF 16, ORF 17, ORF 18, ORF 19, ORF 20, ORF 21, ORF 22, ORF 23, ORF 24, ORF 25, ORF 26, ORF 27, ORF 28, ORF 29, ORF 30, ORF 31, ORF 32, ORF 33, ORF 34, ORF 35, ORF 36, ORF 37, ORF 38, ORF 39, ORF 40, ORF 41, and ORF 42.

15 9. The isolated nucleic acid of claim 4, wherein said nucleic acid comprises a nucleic acid that is a single nucleotide polymorphism (SNP) of a nucleic acid selected from the group consisting of ORF -7, ORF -6, ORF -5, ORF -4, ORF -3, ORF -2, ORF -1, ORF 0, ORF 1, ORF 2, ORF 3, ORF 4, ORF 5, ORF 6, ORF 7, ORF 8, ORF 9, ORF 10, ORF 11, ORF 12, ORF 13, ORF 14, ORF 15, ORF 16, ORF 17, ORF 18, ORF 19, ORF 20, ORF 21, ORF 22, ORF 23, ORF 24, ORF 25, ORF 26, ORF 27, ORF 28, ORF 29, ORF 30, ORF 31, ORF 32, ORF 33, ORF 34, ORF 35, ORF 36, ORF 37, ORF 38, ORF 39, ORF 40, ORF 41, and ORF 42.

20 10. An isolated gene cluster comprising open reading frames encoding polypeptides sufficient to direct the assembly of a C-1027 enediyne or a C-1027 enediyne analogue.

25 11. The gene cluster of claim 10, wherein said gene cluster is present in a bacterium.

12. The gene cluster of claim 11, wherein said gene cluster is present in a bacterium selected from the group consisting of *Actinomycetes*, *Actinoplanetes*, *Actinomadura*, *Micromonospora*, and *Streptomyces*.

22. The host cell of claim 21, wherein said bacterium is selected from the group consisting of Actinomycetes, *Actinoplanetes*, *Actinomadura*, *Micromonospora*, and *Streptomyces*.

5 23. The host cell of claim 21, wherein said bacterium is selected from the group consisting of *Streptomyces globisporus*, *Streptomyces lividans*, *Streptomyces coelicolor*, *Micromonospora echinospira* spp. *calichenisis*, *Actinomadura verrucosopora*, *Micromonospora chersina*, *Streptomyces carzinostaticus*, and *Actinomycete* L585-6.

10 *Sub D2* 24. A method of chemically modifying a biological molecule, said method comprising contacting a biological molecule that is a substrate for a polypeptide encoded by a C-1027 biosynthesis gene cluster open reading frame, with a polypeptide encoded by a C-1027 biosynthesis gene cluster open reading frame whereby said polypeptide chemically modifies said biological molecule.

15 25. The method of claim 24, wherein said polypeptide is an enzyme selected from the group consisting of a hydroxylase, a homocysteine synthase, a dNDP-glucose dehydrogenase, a citrate carrier protein, a C-methyl transferase, an N-methyl transferase, an aminotransferase, a CagA apoprotein, an NDP-glucose synthase, an epimerase, an acyl transferase, a coenzyme F390 synthase, and epoxidase hydrolase, an anthranilate synthase, a glycosyl transferase, a monooxygenase, a type II condensation protein, an aminomutase, a type II adenylation protein, an O-methyl transferase, a P-450
20 hydroxylase, an oxidoreductase, and a proline oxidase.

Sub D3 26. The method of claim 24, wherein said method comprising contacting said biological molecule with at least two different polypeptides encoded by C-1027 biosynthesis gene cluster open reading frames.

25 27. The method of claim 24, wherein said method comprising contacting said biological molecule with at least three different polypeptides encoded by C-1027 biosynthesis gene cluster open reading frames.

28. The method of claim 24, wherein said contacting is in a host cell.

29. The method of claim 28, wherein said host cell is a bacterium.

See D4

30. The method of claim 24, wherein said contacting *ex vivo*.

31. The method of claim 28, wherein said biological molecule is an endogenous metabolite produced by said host cell.

5 See D5

32. The method of claim 28, wherein said biological molecule is an exogenous supplied metabolite.

33. The method of claim 28, wherein said host cell is a eukaryotic cell.

34. The method of claim 33, wherein said eukaryotic cell is selected from the group consisting of a mammalian cell, a yeast cell, a plant cell, a fungal cell, and an insect cell.

10 35. The method of claim 28, wherein said host cell synthesizes sugars and glycosylates the biological molecule.

36. The method of claim 35, wherein said host cell synthesizes deoxysugars.

15 37. The method of claim 24, wherein said method further comprises contacting said biological molecule with a polyketide synthase or a non-ribosomal polypeptide synthetase.

38. The method of claim of claim 24, wherein said contacting is in a bacterial cell.

39. The method of claim of claim 24, wherein said contacting is *ex vivo*.

20 See D6

40. The method of claim 24, wherein said method comprises contacting said biological molecule with at substantially all of the polypeptides encoded by C-1027 biosynthesis gene cluster open reading frames and said method produces an enediynes or enediynes analogue.

25 41. The method of claim 24, wherein said biological molecule is a fatty acid and said biological molecule is contacted with a C-1027 orf polypeptide selected from the

group consisting of an epoxide hydrase, a monooxygenase, an iron-sulfur flavoprotein, a p-450 hydroxylase, an oxidoreductase, and a proline oxidase.

Sub D7

42. The method of claim 41, wherein said biological molecule is a fatty acid and said biological molecule is contacted with a plurality of C-1027 orf polypeptides comprising an epoxide hydrase, a monooxygenase, an iron-sulfur flavoprotein, a p-450 hydroxylase, an oxidoreductase, and a proline oxidase.

43. The method of claim 42, wherein said biological molecule is contacted with polypeptides encoded by ORF17, ORF20, ORF21, ORF29, ORF30, ORF32, ORF35, and ORF38.

44. The method of claim 41, wherein said biological molecule is contacted with polypeptides encoded by ORF 15, ORF 16, ORF 28, ORF3, ORF 14, and ORF 13.

45. The method of claim 44 wherein said biological molecule is also contacted with polypeptides encoded by ORF 4 and ORF 3.

46. The method of claim 24, wherein said method comprises contacting a sugar with one or more C-1027 open reading frame polypeptides selected from the group consisting of a dNDP-glucose synthase, a dNDP glucose dehydratase, an epimerase, an aminotransferase, a C-methyltransferase, an N-methyltransferase, and a glycosyl transferase.

47. The method of claim 46, wherein said method comprises contacting a dNDP-glucose with a plurality of C-1027 open reading frame polypeptides comprising a dNDP-glucose synthase, a dNDP glucose dehydratase, an epimerase, an aminotransferase, a C-methyltransferase, an N-methyltransferase, and a glycosyl transferase.

48. The method of claim 24, wherein said method comprises contacting an amino acid with one or one or more C-1027 open reading frame polypeptides selected from the group consisting of a hydroxylase, an aminomutase, a type II NRPS condensation enzyme, a type II NRPS adenylation enzyme, and a type II peptidyl carrier protein.

49. The method of claim 48, wherein said method comprises contacting an amino acid with a plurality of C-1027 open reading frame polypeptides comprising a

hydroxylase, a halogenase, an aminomutase, a type II NRPS condensation enzyme, a type II NRPS adenylation enzyme, and a type II peptidyl carrier protein.

50. The method of claim 48, wherein said amino acid is a tyrosine.

51. A method of synthesizing a chromaprotein type enediyne core, said
5 method comprising contacting a fatty acid with one or more C-1027 orf polypeptides
selected from the group consisting of an epoxide hydrase, a monooxygenase, an iron-sulfur
flavoprotein, a p-450 hydroxylase, an oxidoreductase, and a proline oxidase.

52. The method of claim 51, wherein said fatty acid is contacted with a
plurality of C-1027 orf polypeptides comprising an epoxide hydrase, a monooxygenase, an
10 iron-sulfur flavoprotein, a p-450 hydroxylase, an oxidoreductase, and a proline oxidase.

53. The method of claim 52, wherein said fatty acid is contacted with
polypeptides encoded by ORF17, ORF20, ORF21, ORF29, ORF30, ORF32, ORF35, and
ORF38.

54. A method of synthesizing a deoxysugar, said method comprising
15 contacting a sugar with one or more C-1027 open reading frame polypeptides selected from
the group consisting of a dNDP-glucose synthase, a dNDP glucose dehydratase, an
epimerase, an aminotransferase, a C-methyltransferase, an N-methyltransferase, and a
glycosyl transferase.

55. The method of claim 54, wherein said method comprises contacting a
20 dNDP-glucose with a plurality of C-1027 open reading frame polypeptides comprising a
dNDP-glucose synthase, a dNDP glucose dehydratase, an epimerase, an aminotransferase, a
C-methyltransferase, an N-methyltransferase, and a glycosyl transferase.

56. The method of claim 55, wherein said dNDP-glucose is contacted with
polypeptides encoded by ORF17, ORF20, ORF21, ORF29, ORF30, ORF32, ORF35, and
25 ORF38.

57. A method of synthesizing a beta amino acid, said method comprising
contacting an amino acid with one or one or more C-1027 open reading frame polypeptides
selected from the group consisting of a hydroxylase, an aminomutase, a type II NRPS

condensation enzyme, a type II NRPS adenylation enzyme, and a type II peptidyl carrier protein.

58. The method of claim 57, wherein said method comprises contacting an amino acid with a plurality of C-1027 open reading frame polypeptides comprising a hydroxylase, a halogenase, an aminomutase, a type II NRPS condensation enzyme, a type II NRPS adenylation enzyme, and a type II peptidyl carrier protein.

59. The method of claim wherein said amino acid is contacted with polypeptides encoded by ORF4, ORF11, ORF24, ORF23, ORF25, and ORF26.

60. The method of claim 57, wherein said amino acid is a tyrosine.

61. A method of synthesizing an enediyne or an enediyne analogue said method comprising:

culturing a cell comprising a recombinantly modified C-1027 gene cluster under conditions whereby said cell expresses said enediyne or enediyne analogue; and

recovering said enediyne or enediyne analogue.

62. The method of claim 61, wherein said gene cluster is present in a bacterium.

63. The gene cluster of claim 62, wherein said gene cluster is present in a bacterium selected from the group consisting of Actinomycetes, *Actinoplanetes*, *Actinomadura*, *Micromonospora*, and *Streptomyces*.

64. The gene cluster of claim 62, wherein said gene cluster is present in a bacterium selected from the group consisting *Streptomyces globisporus*, *Streptomyces lividans*, *Streptomyces coelicolor*, *Micromonospora echinospora* spp. *calichenisis*, *Actinomadura verrucosopora*, *Micromonospora chersina*, *Streptomyces carzinostaticus*, and *Actinomycete* L585-6.

65. The method of claim 61, wherein said gene cluster is present in a eukaryotic cell.

